AD-A060 231 NAVY UNDERWATER SOUND LAB NEW LONDON CONN F/G 17/1

AN ALGORITHM FOR THE DIRECT ESTIMATION OF INVERSE COVARIANCE MA--ETC(U)

FEB 69 N L OWSLEY

UNCLASSIFIED USL-TM-2242-32-69 NL

| OF |

AD
A060231

END
DATE
FILMED
12-78
DDC

Code No. 1
Copy No. 41

U. S. NAVY UNDERWATER SOUND LABORATORY
FORT TRUMBULL, NEW LONDON, CONNECTICUT

C
NW

2242-32-69

6

AN ALGORITHM FOR THE DIRECT ESTIMATION OF INVERSE COVARIANCE MATRICES

by

10 Norman L./Owsley

USL Technical Memorandum No. 2242-32-69

11 11 February 1969

Introduction

The design of many types of optimum digital signal processing schemes involves an assumed knowledge of various types of covariance matrices. If these matrix quantities are unknown a priori they must be estimated. In addition, many processor design criteria require a knowledge of inverse covariance matrices for the purpose of implementing various digital noise "whitening" operations. Generally, the method for obtaining an estimated inverse covariance matrix is to estimate the original matrix and then invert it digitally. If the dimensionality of the covariance matrix doesn't preclude a digital inversion, then, in many environments, the time consumed by the inversion process does. This memorandum derives an algorithm for directly estimating the inverse of a covariance matrix. The estimation technique used is that of multidimensional gradient search. The method is applicable in a nonstationary noise environment with the inverse of an arbitrary positive definite matrix required as an initial condition.

9 Technical memo. 12 10p.

Inverse Covariance Matrices and Residues

Consider the finite dimensional sample vector $\underline{x}$. If this vector is obtained from a sample function of a random process x, then a positive definite covariance matrix

14 USL-TM-2242-32-69

Encl (1) to USL Ser: 2242-44 of

$$C = Q^{-1}$$

$$= E\left\{\underline{x}\,\underline{x}^T\right\}$$

$$= \begin{bmatrix} c_{11} & c_{12} & c_{13} & \cdots & c_{1K} \\ c_{21} & c_{22} & c_{23} & \cdots & c_{2K} \\ c_{31} & c_{32} & c_{33} & \cdots & c_{3K} \\ \vdots & & & \ddots & \vdots \\ c_{K1} & c_{K2} & c_{K3} & \cdots & c_{KK} \end{bmatrix}$$

can be defined. The notation $E\{\ \}$ indicates the expectation operation and a superscript "T" specifies a vector transpose. It is assumed that the sample vector $\underline{x}$ has zero mean. That is,

$$E\{\underline{x}\} = \underline{0}$$

with the implication that x is a zero mean random process. Furthermore, the inverse covariance matrix Q is specified with the notation

$$Q = \begin{bmatrix} q_{11} & q_{12} & q_{13} & \cdots & q_{1K} \\ q_{21} & q_{22} & q_{23} & \cdots & q_{2K} \\ q_{31} & q_{32} & q_{33} & \cdots & q_{3K} \\ \vdots & & & \ddots & \vdots \\ q_{K1} & q_{K2} & q_{K3} & \cdots & q_{KK} \end{bmatrix}.$$

The vector

$$\underline{v} = Q\,\underline{x}$$

is now defined by

$$\underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_K \end{bmatrix}$$

where

$$v_k = \sum_{\ell=1}^{K} q_{k\ell} x_\ell \quad .$$

If we further define the vectors $\underline{Q}_k$ and $\underline{x}$ as

$$\underline{Q}_k = \begin{bmatrix} q_{k1} \\ q_{k2} \\ \vdots \\ q_{k\ell} \\ \vdots \\ q_{kK} \end{bmatrix}$$

and

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\ell \\ \vdots \\ x_K \end{bmatrix}$$

respectively, then an element in the $\underline{V}$ vector has the value

$$v_k = \underline{Q}_k^T \underline{x}$$

$$= \frac{\text{residue of } x_k = r_k}{\text{variance of } r_k} \,.$$

The residue of the sample x_k is now said to be given by

$$r_k = x_k - \hat{x}_k$$

where $\hat{x}_k$ is the "best" linear mean square (LMS) estimate[1] of x_k in terms of the other (K - 1) terms of $\underline{x}$. To be more specific,

$$\hat{x}_k = \sum_{\ell \neq k} \beta_{k\ell}\, x_\ell$$

where the regression coefficients, $\beta_{k\ell}$ have been selected to minimize

$$E\{ r_k^2 \} .$$

<u>Gradient Search Method for Obtaining the Regression Coefficients</u>

At the ith step i[illegible] a sequence of observations the LMS residue for the observation $x_k(i)$ is

$$r_k(i) = x_k(i) - \hat{x}_k(i)$$

$$= x_k(i) - \sum_{\ell = k} \beta_{k\ell}(i)\, x_\ell(i) .$$

Therefore,

$$r_k^2(i) = \left[x_k(i) - \sum_{\ell = k} \beta_{k\ell}(i)\, x_\ell(i) \right]^2$$

and the gradient (or partial derivative) of $r_k^2(i)$ with respect to the regression coefficient $\beta_{kj}(i)$ is

$$\frac{\partial r_k^2(i)}{\partial \beta_{kj}(i)} = 2 \left[x_k(i) - \sum_{\ell \neq k} \beta_{k\ell}(i)\, x_\ell(i) \right] \left[- x_{j \neq k}(i) \right]$$

$$= - 2\, r_k(i)\, x_{j \neq k}(i) .$$

Therefore,

$$E\left\{\frac{\partial r_k^2(i)}{\partial \beta_{kj}(i)}\right\} = -2\left[c_{kj}(i) - \sum_{l \neq k} \beta_{kl}(i)\, c_{lj}(i)\right]_{j \neq k}$$

is the expected value of the elemental change in the residue variance $E\{r_k^2(i)\}$ with respect to the regression coefficient $\beta_{kj}(i)$. The recursive formula for a modified gradient search technique can now be specified as

$$\beta_{kj}(i+1) = \beta_{kj}(i) + K_r \frac{\partial r_k^2(i)}{\partial \beta_{kj}(i)}$$

$$= \beta_{kj}(i) - 2K_r\, r_k(i)\, x_{j \neq k}(i),$$

where

$$\frac{\partial r_k^2(i)}{\partial \beta_{kj}(i)}$$

is used as a "noisy" estimate of

$$E\left\{\frac{\partial r_k^2(i)}{\partial \beta_{kj}(i)}\right\}.$$

In the above, $\beta_{kj}(i+1)$ is an updated estimate of β_{kj} in terms of the old estimate $\beta_{kj}(i)$ and the ith observation.

In vector form,

$$\underline{\beta}_k(i+1) = \underline{\beta}_k(i) - 2K_r\, r_k(i)\, \underline{x}_k(i)$$

where K_r is a negative scalar constant controlling the rate of convergence and stability for the estimator vector $\underline{\beta}_k(i+1)$ and

$$\underline{x}_k(i) = \begin{bmatrix} x_1(i) \\ x_2(i) \\ \vdots \\ x_{k-1}(i) \\ x_{k+1}(i) \\ \vdots \\ x_K(i) \end{bmatrix} .$$

The properties of this type of estimator are given by Widrow et al.[2]

Relation Between the Regression Coefficient Vector $\underline{\beta}_k(i)$ and $\underline{Q}_k(i)$

Recall that

$$v_k = \underline{Q}_k^T \underline{x} ,$$

which can also be written as

$$v_k = \frac{x_k - \sum_{\ell \neq k} \beta_{k\ell} x_\ell}{E\{r_k^2\}} .$$

Now define an augmented regression coefficient vector for the kth residue as

$$\underline{B}_k = \begin{bmatrix} -\beta_{k1} \\ -\beta_{k2} \\ \vdots \\ 1 \\ \vdots \\ -\beta_{kK} \end{bmatrix} \begin{array}{l} \\ \\ \\ \longleftarrow \text{kth element .} \\ \\ \\ \end{array}$$

The element v_k now becomes

$$v_k = \frac{\underline{B}_k^T \underline{x}}{E\{r_k^2\}} .$$

We now identify $\underline{Q}_k$ with $\underline{B}_k$ by the relationship

$$\underline{Q}_k = \frac{1}{E\{r_k^2\}} \underline{B}_k$$

which for the ith observation can be approximated by

$$\hat{\underline{Q}}_k(i) = \frac{1}{\langle r_k^2(i) \rangle} \underline{B}_k(i) .$$

The "normalizing" term, $\langle r_k^2(i) \rangle$, can easily be obtained from a finite time digital averaging process. Finally, the estimated inverse correlation matrix at the ith step of an observation sequence can be written as

$$\hat{Q}(i) = \begin{bmatrix} \hat{\underline{Q}}_1^T(i) \\ \hat{\underline{Q}}_2^T(i) \\ \vdots \\ \hat{\underline{Q}}_k^T(i) \\ \vdots \\ \hat{\underline{Q}}_K^T(i) \end{bmatrix} .$$

Conclusion

The algorithm presented in this memorandum has direct application to the realistic implementation of various types of signal processors based on an adaptive prewhitening approach. This technique relates directly to the area of transient signal detection and can be extended to systems employing spatially distributed receiver elements.[3] In addition, adaptive beamforming techniques require the knowledge of inverted covariance matrices to specify an optimal filter configuration.[2] The direct estimation of inverse spectral covariance matrices would also expedite the implementation of signal processing schemes which are frequency domain oriented.[4] When considered in the digital signal processing context, the advantage of the technique presented

herein is that it completely avoids the issue of digital matrix inversion and still allows the inherent advantage of noise whitening.

Norman L. Owsley

NORMAN L. OWSLEY

References

1. Cramer, H., "Mathematical Methods of Statistics", pp. 302-305, Princeton Univ. Press, 1951.

2. Widrow, B., Mantey, P., Griffiths, L., and Goode, B., "Adaptive Antenna Systems", Proceedings of the IEEE, Vol. 55, N. 12, December 1967.

3. Bryn, F., "Optimum Theoretical Structures of Sonar Systems Employing Spatially-Distributed Receiving Elements", SACLANT ASW Research Center, September 1968.

4. Owsley, N., "The Parametric Sequential Classification of Spectral Patterns with Application to Signal Detection and Extraction", Proceedings of the 1968 IEEE EASCON, September 1968.